

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-067397

(43)Date of publication of application : 07.03.2003

(51)Int.Cl. G06F 17/30
G06F 12/00

(21)Application number : 2002-167767

(71)Applicant : MATSUSHITA ELECTRIC IND CO
LTD

(22)Date of filing : 07.06.2002

(72)Inventor : SHIMOJIMA TAKASHI
NOGUCHI NAOHIKO

(30)Priority

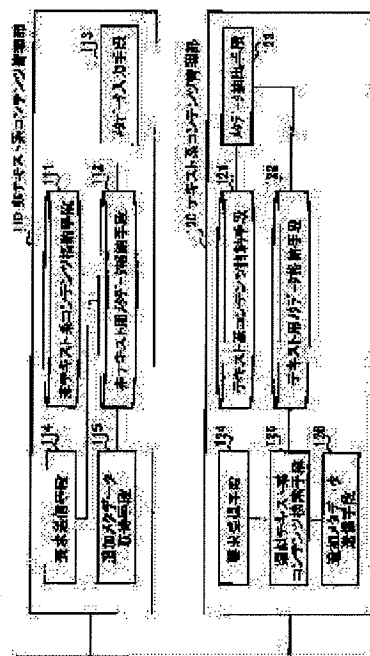
Priority number : 2001175136 Priority date : 11.06.2001 Priority country : JP

(54) CONTENT CONTROL SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To conduct processing for imparting a meta-data as to a meaning content with reduced labor, in a multimedia content such as an image and a voice.

SOLUTION: An original meta-data imparted to a non-text content by a hand of a human being is transmitted from a non-text content control part 110 to a text content control part 120, a similar text content is retrieved by a similar text content retrieving means 125 in the text content control part 120, using the original meta-data, and the meta-data extracted automatically in the similar text content is transmitted to the non-text content control part 110 as an additional meta-data to the non-text content.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-67397

(P2003-67397A)

(43) 公開日 平成15年3月7日(2003.3.7)

(51) Int.Cl. ⁷	識別記号	F I	テームト* (参考)
G 0 6 F 17/30	2 1 0	G 0 6 F 17/30	2 1 0 C 5 B 0 7 5
	1 7 0		1 7 0 A 5 B 0 8 2
			1 7 0 G
	3 5 0		3 5 0 C
12/00	5 2 0	12/00	5 2 0 E
審査請求 未請求 請求項の数29 O L (全 21 頁)			

(21) 出願番号 特願2002-167767(P2002-167767)
(22) 出願日 平成14年6月7日(2002.6.7)
(31) 優先権主張番号 特願2001-175136(P2001-175136)
(32) 優先日 平成13年6月11日(2001.6.11)
(33) 優先権主張国 日本 (J P)

(71) 出願人 000005821
松下電器産業株式会社
大阪府門真市大字門真1006番地
(72) 発明者 下島 崇
大阪府門真市大字門真1006番地 松下電器
産業株式会社内
(72) 発明者 野口 直彦
大阪府門真市大字門真1006番地 松下電器
産業株式会社内
(74) 代理人 100105050
弁理士 鷲田 公一

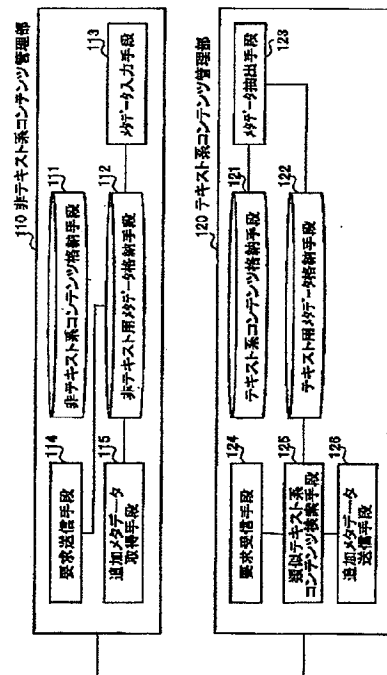
最終頁に続く

(54) 【発明の名称】 コンテンツ管理システム

(57) 【要約】

【課題】 映像や音声といったマルチメディアコンテンツに対して、その意味内容に関するメタデータを付与する処理をより少ない労力で実現することを目的とする。

【解決手段】 非テキスト系コンテンツ管理部110から、テキスト系コンテンツ管理部120に対して非テキスト系コンテンツに人間の手によって付与されたオリジナルメタデータを送信し、テキスト系コンテンツ管理部120の類似テキスト系コンテンツ検索手段125にて前記オリジナルメタデータを利用して類似するテキスト系コンテンツを検索し、当該類似テキスト系コンテンツに対して自動抽出されたメタデータを、非テキスト系コンテンツ管理部110に対して送信する。



【特許請求の範囲】

【請求項1】 メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を送信する送信手段と、

追加メタデータを受信する受信手段と、

受信された追加メタデータをメタデータ追加処理対象である非テキスト系コンテンツに付与する付与手段と、を有することを特徴とする非テキスト系コンテンツ管理装置。

【請求項2】 前記付与手段は、

受信された追加メタデータをそのままメタデータ追加処理対象である非テキスト系コンテンツに付与する、ことを特徴とする請求項1記載の非テキスト系コンテンツ管理装置。

【請求項3】 前記付与手段は、

受信された追加メタデータを、オリジナルメタデータと重複する部分を除いてメタデータ追加処理対象である非テキスト系コンテンツに付与する、ことを特徴とする請求項1記載の非テキスト系コンテンツ管理装置。

【請求項4】 メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を受信する受信手段と、

受信された追加メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索する検索手段と、

検索されたテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得する取得手段と、

取得された追加メタデータを送信する送信手段と、を有することを特徴とするテキスト系コンテンツ管理装置。

【請求項5】 前記取得手段は、

類似するテキスト系コンテンツが複数検索された場合、検索された複数の類似するテキスト系コンテンツのうち類似度が最も高いテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得する、ことを特徴とする請求項4記載のテキスト系コンテンツ管理装置。

【請求項6】 前記取得手段は、

類似するテキスト系コンテンツが複数検索された場合、検索された複数の類似するテキスト系コンテンツのうち類似度が高い方から複数の所定数のテキスト系コンテンツにそれぞれあらかじめ付与されたメタデータの組を追加メタデータとして取得する、

ことを特徴とする請求項4記載のテキスト系コンテンツ管理装置。

【請求項7】 非テキスト系コンテンツを扱う非テキスト系コンテンツ管理装置と、テキスト系コンテンツを扱うテキスト系コンテンツ管理装置とからなるコンテンツ管理システムであって、

前記非テキスト系コンテンツ管理装置は、

メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を前記テキスト系コンテンツ管理装置へ送信する第1送信手段と、

10 前記テキスト系コンテンツ管理装置から追加メタデータを受信する第2受信手段と、

受信された追加メタデータをメタデータ追加処理対象である非テキスト系コンテンツに付与する付与手段と、を有し、

前記テキスト系コンテンツ管理装置は、

前記非テキスト系コンテンツ管理装置から追加メタデータ要求を受信する第2受信手段と、

受信された追加メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索する検索手段と、

検索されたテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得する取得手段と、

取得された追加メタデータを前記非テキスト系コンテンツ管理装置へ送信する第2送信手段と、

を有することを特徴とするコンテンツ管理システム。

【請求項8】 前記非テキスト系コンテンツ管理装置および前記テキスト系コンテンツ管理装置は、同一のコンピュータ上に存在することを特徴とする請求項7記載のコンテンツ管理システム。

【請求項9】 前記非テキスト系コンテンツ管理装置および前記テキスト系コンテンツ管理装置は、異なるコンピュータ上に存在し、情報伝送可能に接続されていることを特徴とする請求項7記載のコンテンツ管理システム。

【請求項10】 関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を送信する送信手段と、

関連コンテンツ付属メタデータを受信する受信手段と、受信された関連コンテンツ付属メタデータを元に、関連コンテンツ情報作成処理対象である非テキスト系コンテンツに対する関連コンテンツ情報を作成する作成手段と、

を有することを特徴とする非テキスト系コンテンツ管理装置。

【請求項11】 前記関連コンテンツ情報は、関連コンテンツ情報作成処理対象である非テキスト系コンテンツと類似する非テキスト系コンテンツにあらかじめ付与さ

れたコンテンツIDを含み、
前記作成手段は、
受信された関連コンテンツ付属メタデータを元に、関連
コンテンツ情報作成処理対象である非テキスト系コンテ
ンツと類似する非テキスト系コンテンツを検索する検索
手段と、
検索された非テキスト系コンテンツにあらかじめ付与さ
れたコンテンツIDを取得する取得手段と、
を有することを特徴とする請求項10記載の非テキスト
系コンテンツ管理装置。

【請求項12】 前記取得手段は、
類似する非テキスト系コンテンツが複数検索された場
合、検索された複数の類似する非テキスト系コンテ
ンツのうち類似度が最も高い非テキスト系コンテンツにあら
かじめ付与されたコンテンツIDを取得する、
ことを特徴とする請求項11記載の非テキスト系コンテ
ンツ管理装置。

【請求項13】 前記取得手段は、
類似する非テキスト系コンテンツが複数検索された場
合、検索された複数の類似する非テキスト系コンテ
ンツのうち類似度が高い方から複数の所定数の非テキスト系
コンテンツにそれぞれあらかじめ付与されたコンテンツ
IDの組を取得する、
ことを特徴とする請求項11記載の非テキスト系コンテ
ンツ管理装置。

【請求項14】 関連コンテンツ情報作成処理対象であ
る非テキスト系コンテンツにあらかじめ付与されたオリ
ジナルメタデータを含む関連コンテンツ付属メタデータ
要求を受信する受信手段と、
受信された関連コンテンツ付属メタデータ要求に含まれ
るオリジナルメタデータを元に、当該オリジナルメタデ
ータに対応する非テキスト系コンテンツと類似するテキ
スト系コンテンツを検索する検索手段と、
検索されたテキスト系コンテンツと関連するテキスト系
コンテンツにあらかじめ付与されたメタデータを関連コ
ンテンツ付属メタデータとして取得する取得手段と、
取得された関連コンテンツ付属メタデータを送信する送
信手段と、
を有することを特徴とするテキスト系コンテンツ管理装
置。

【請求項15】 前記取得手段は、
類似するテキスト系コンテンツが複数検索された場合、
検索された複数の類似するテキスト系コンテンツのうち
類似度が最も高いテキスト系コンテンツにあらかじめ付
与されたメタデータを関連コンテンツ付属メタデータと
して取得する、
ことを特徴とする請求項14記載のテキスト系コンテ
ンツ管理装置。

【請求項16】 前記取得手段は、
類似するテキスト系コンテンツが複数検索された場合、

検索された複数の類似するテキスト系コンテンツのうち
類似度が高い方から複数の所定数のテキスト系コンテ
ンツにそれぞれあらかじめ付与されたメタデータの組を関
連コンテンツ付属メタデータとして取得する、
ことを特徴とする請求項14記載のテキスト系コンテ
ンツ管理装置。

【請求項17】 非テキスト系コンテンツを扱う非テキ
スト系コンテンツ管理装置と、テキスト系コンテンツを
扱うテキスト系コンテンツ管理装置とからなるコンテ
ンツ管理システムであって、

前記非テキスト系コンテンツ管理装置は、
関連コンテンツ情報作成処理対象である非テキスト系コ
ンテンツにあらかじめ付与されたオリジナルメタデータ
を含む関連コンテンツ付属メタデータ要求を前記テキス
ト系コンテンツ管理装置へ送信する第1送信手段と、
前記テキスト系コンテンツ管理装置から関連コンテンツ
付属メタデータを受信する第1受信手段と、
受信された関連コンテンツ付属メタデータを元に、関連
コンテンツ情報作成処理対象である非テキスト系コンテ
ンツに対する関連コンテンツ情報を作成する作成手段
と、を有し、

前記テキスト系コンテンツ管理装置は、
前記非テキスト系コンテンツ管理装置から関連コンテ
ンツ付属メタデータ要求を受信する第2受信手段と、
受信された関連コンテンツ付属メタデータ要求に含まれ
るオリジナルメタデータを元に、当該オリジナルメタデ
ータに対応する非テキスト系コンテンツと類似するテキ
スト系コンテンツを検索する検索手段と、
検索されたテキスト系コンテンツと関連するテキスト系
コンテンツにあらかじめ付与されたメタデータを関連コ
ンテンツ付属メタデータとして取得する取得手段と、
取得された関連コンテンツ付属メタデータを前記非テキ
スト系コンテンツ管理装置へ送信する第2送信手段と、
を有することを特徴とするコンテンツ管理システム。

【請求項18】 前記非テキスト系コンテンツ管理装置
および前記テキスト系コンテンツ管理装置は、同一のコ
ンピュータ上に存在することを特徴とする請求項17記
載のコンテンツ管理システム。

【請求項19】 前記非テキスト系コンテンツ管理装置
および前記テキスト系コンテンツ管理装置は、異なるコ
ンピュータ上に存在し、情報伝送可能に接続されている
ことを特徴とする請求項17記載のコンテンツ管理シ
ステム。

【請求項20】 一の非テキスト系コンテンツと類似す
るテキスト系コンテンツを利用して、前記一の非テキ
スト系コンテンツと関連する他の非テキスト系コンテ
ンツを検索することを特徴とするコンテンツ管理装置。

【請求項21】 前記検索は、前記一の非テキスト系コ
ンテンツと類似するテキスト系コンテンツとさらに関連
するテキスト系コンテンツを利用して行なわれることを

特徴とする請求項20記載のコンテンツ管理装置。

【請求項22】 非テキスト系コンテンツを扱う非テキスト系コンテンツ管理装置と、テキスト系コンテンツを扱うテキスト系コンテンツ管理装置とからなるコンテンツ管理システムにおけるメタデータ追加方法であって、前記非テキスト系コンテンツ管理装置が、メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を前記テキスト系コンテンツ管理装置へ送信するステップと、

前記テキスト系コンテンツ管理装置が、前記非テキスト系コンテンツ管理装置から追加メタデータ要求を受信するステップと、

前記テキスト系コンテンツ管理装置が、受信した追加メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、

前記テキスト系コンテンツ管理装置が、検索したテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得するステップと、

前記テキスト系コンテンツ管理装置が、取得した追加メタデータを前記非テキスト系コンテンツ管理装置へ送信するステップと、

前記非テキスト系コンテンツ管理装置が、前記テキスト系コンテンツ管理装置から追加メタデータを受信するステップと、

前記非テキスト系コンテンツ管理装置が、受信した追加メタデータをメタデータ追加処理対象である非テキスト系コンテンツに付与するステップと、

を有することを特徴とするメタデータ追加方法。

【請求項23】 非テキスト系コンテンツを扱う非テキスト系コンテンツ管理装置と、テキスト系コンテンツを扱うテキスト系コンテンツ管理装置とからなるコンテンツ管理システムにおける関連コンテンツ情報作成方法であって、

前記非テキスト系コンテンツ管理装置が、関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を前記テキスト系コンテンツ管理装置へ送信するステップと、

前記テキスト系コンテンツ管理装置が、前記非テキスト系コンテンツ管理装置から関連コンテンツ付属メタデータ要求を受信するステップと、

前記テキスト系コンテンツ管理装置が、受信した関連コンテンツ付属メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、

前記テキスト系コンテンツ管理装置が、検索したテキス

ト系コンテンツと関連するテキスト系コンテンツにあらかじめ付与されたメタデータを関連コンテンツ付属メタデータとして取得するステップと、

前記テキスト系コンテンツ管理装置が、取得した関連コンテンツ付属メタデータを前記非テキスト系コンテンツ管理装置へ送信するステップと、

前記非テキスト系コンテンツ管理装置が、前記テキスト系コンテンツ管理装置から関連コンテンツ付属メタデータを受信するステップと、

10 前記非テキスト系コンテンツ管理装置が、受信した関連コンテンツ付属メタデータを元に、関連コンテンツ情報作成処理対象である非テキスト系コンテンツに対する関連コンテンツ情報を作成するステップと、
を有することを特徴とする関連コンテンツ情報作成方法。

【請求項24】 メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を送信するステップと、

20 追加メタデータを受信するステップと、

受信された追加メタデータをメタデータ追加処理対象である非テキスト系コンテンツに付与するステップと、
をコンピュータに実行させるためのコンテンツ管理プログラム。

【請求項25】 メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を受信するステップと、

30 受信された追加メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、

検索されたテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得するステップと、

取得された追加メタデータを送信するステップと、
をコンピュータに実行させるためのコンテンツ管理プログラム。

【請求項26】 コンピュータを、非テキスト系コンテンツを扱う非テキスト系コンテンツ管理部およびテキスト系コンテンツを扱うテキスト系コンテンツ管理部として機能させるためのコンテンツ管理プログラムであって、前記コンテンツ管理プログラムは、

前記非テキスト系コンテンツ管理部が、メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を前記テキスト系コンテンツ管理部へ送信するステップと、

50 前記テキスト系コンテンツ管理部が、前記非テキスト系コンテンツ管理部から追加メタデータ要求を受信するス

トップと、
 前記テキスト系コンテンツ管理部が、受信した追加メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、
 前記テキスト系コンテンツ管理部が、検索したテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得するステップと、
 前記テキスト系コンテンツ管理部が、取得した追加メタデータを前記非テキスト系コンテンツ管理部へ送信するステップと、
 前記非テキスト系コンテンツ管理部が、前記テキスト系コンテンツ管理部から追加メタデータを受信するステップと、
 前記非テキスト系コンテンツ管理部が、受信した追加メタデータをメタデータ追加処理対象である非テキスト系コンテンツに付与するステップと、
 を有することを特徴とするコンテンツ管理プログラム。
 【請求項27】 関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を送信するステップと、
 関連コンテンツ付属メタデータを受信するステップと、
 受信された関連コンテンツ付属メタデータを元に、関連コンテンツ情報作成処理対象である非テキスト系コンテンツに対する関連コンテンツ情報を作成するステップと、
 をコンピュータに実行させるためのコンテンツ管理プログラム。
 【請求項28】 関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を受信するステップと、
 受信された関連コンテンツ付属メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、
 検索されたテキスト系コンテンツと関連するテキスト系コンテンツにあらかじめ付与されたメタデータを関連コンテンツ付属メタデータとして取得するステップと、
 取得された関連コンテンツ付属メタデータを送信するステップと、
 をコンピュータに実行させるためのコンテンツ管理プログラム。
 【請求項29】 コンピュータを、非テキスト系コンテンツを扱う非テキスト系コンテンツ管理部およびテキスト系コンテンツを扱うテキスト系コンテンツ管理部として機能させるためのコンテンツ管理プログラムであって、前記コンテンツ管理プログラムは、

前記非テキスト系コンテンツ管理部が、関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を前記テキスト系コンテンツ管理部へ送信するステップと、
 前記テキスト系コンテンツ管理部が、前記非テキスト系コンテンツ管理部から関連コンテンツ付属メタデータ要求を受信するステップと、
 前記テキスト系コンテンツ管理部が、受信した関連コンテンツ付属メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、
 前記テキスト系コンテンツ管理部が、検索したテキスト系コンテンツと関連するテキスト系コンテンツにあらかじめ付与されたメタデータを関連コンテンツ付属メタデータとして取得するステップと、
 前記テキスト系コンテンツ管理部が、取得した関連コンテンツ付属メタデータを前記非テキスト系コンテンツ管理部へ送信するステップと、
 前記非テキスト系コンテンツ管理部が、前記テキスト系コンテンツ管理部から関連コンテンツ付属メタデータを受信するステップと、
 前記非テキスト系コンテンツ管理部が、受信した関連コンテンツ付属メタデータを元に、関連コンテンツ情報作成処理対象である非テキスト系コンテンツに対する関連コンテンツ情報を作成するステップと、
 を有することを特徴とするコンテンツ管理プログラム。

【発明の詳細な説明】

30 【0001】

【発明の属する技術分野】本発明は、マルチメディアコンテンツを管理するコンピュータシステムにおいて、特に映像や音声といった非テキスト系コンテンツに対するメタデータの付与に関する。

【0002】

【従来の技術】インターネットの普及に伴い、我々は様々なコンテンツに容易にアクセス可能となった。更に近年では、ADSL (Asymmetric Digital Subscriber Line)、FTTH (Fiber To The Home) といった技術を利用した通信網のブロードバンド (広帯域) 化によって、従来のような比較的データサイズの小さいテキストや画像中心のコンテンツのみならず、映像や音声といったマルチメディア情報を快適に利用可能な環境が整いつつあり、今後更に様々なコンテンツが提供されるようになることが予想される。

【0003】このように、利用可能なコンテンツが膨大になるのに比例して、欲しいコンテンツを見つける検索技術や、不要なコンテンツを削除するフィルタリング技術がより重要になる。特に映像や音声といったマルチメディアコンテンツはテキスト系のコンテンツとは異な

50

り、そのままの形で検索およびフィルタリング対象となり得ない。

【0004】そこで、このような検索やフィルタリングを行なうためには、コンテンツに対してその特徴を記述したメタデータが必要であり、そのための技術が求められる。特にここで、コンテンツの意味的内容を記述したメタデータを考えてみると、テキスト系データの場合には、既に様々な研究がなされている。例えば、アメリカ政府関係の主催するTipsterプロジェクトでは、テキスト処理に関する技術の推進を行なっており、その中でテキストからの情報抽出技術の研究開発も行なわれている(Tipsterプロジェクトについては、福本淳一、関根聡、江里口善生：「MJC-7, Tipster参加報告」, 情報処理学会自然言語処理, 127-14, 1998を参照)。

【0005】一方、映像や音声といった非テキスト系コンテンツに対するメタデータの枠組みの一例として、MPEG-7 (正式名称"Multimedia Content Description Interface", [ISO/IEC 15938]) が挙げられる。MPEG-7は、マルチメディア情報の内容を記述するための記述子を規定する国際標準規格であり、この記述子によってマルチメディアコンテンツの内容に基づく検索やフィルタリングなどを実現することを目的とする。

【0006】

【発明が解決しようとする課題】ところが、MPEG-7がメタデータ付与の対象とする映像や音声といった非テキスト系コンテンツの場合には、例えば、ニュース番組映像のコンテンツのある時間帯について、どのような内容のニュースであるかといったメタデータを自動的に抽出する技術は存在せず、人間の手によってメタデータが付与されているのが現実である。

【0007】しかしながら、このような人間の手作業によるメタデータ付与は非常に時間と労力がかかり非効率な方法であるので、コンテンツプロバイダとしては、コストの面から考えて、それほど豊富なメタデータを非テキスト系コンテンツに対して付与できない。

【0008】また、人間の手で付与したメタデータがそれほど豊富でないために、関連する別のニュース映像コンテンツを検索するといったことを高い精度で行なうことができない。

【0009】本発明は、かかる点に鑑みてなされたものであり、非テキスト系コンテンツに対して最低限人間の手で付与された意味内容に関するメタデータを元に、より豊富なメタデータを付加することができる、また、非テキスト系コンテンツ同士の内容の関連性を導き出すことができるコンテンツ管理システムを提供することを目的とする。

【0010】

【課題を解決するための手段】本発明は、上記目的を達成するため、メタデータ追加処理対象の非テキスト系コンテンツに対して人間の手によって付与されたメタデー

タを利用して、類似するテキスト系コンテンツを検索する類似テキスト系コンテンツ検索手段と、当該類似テキスト系コンテンツに対して自動的に抽出されたメタデータを、メタデータ追加処理対象の非テキスト系コンテンツに対する追加メタデータとして取得する追加メタデータ取得手段を備えることにより、非テキスト系コンテンツに対するメタデータを、人間の手によって付与された、限られた数から増やすことが可能となる。

【0011】また、本発明によれば、関連コンテンツ情報作成処理対象の非テキスト系コンテンツに対して人間の手によって付与されたメタデータを利用して、類似するテキスト系コンテンツを検索する類似テキスト系コンテンツ検索手段と、当該類似テキスト系コンテンツに対して関連付けされたテキスト系コンテンツに対して自動的に抽出されたメタデータ(関連コンテンツ付属メタデータ)を利用して、類似する非テキスト系コンテンツを検索する類似非テキスト系コンテンツ検索手段を備えることにより、最低限人間の手によって付与されたメタデータのみでは導き出すことが出来なかった非テキスト系コンテンツ同士の関連性を導き出すことが可能となる。

【0012】本発明の非テキスト系コンテンツ管理装置は、メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を送信する送信手段と、追加メタデータを受信する受信手段と、受信された追加メタデータをメタデータ追加処理対象である非テキスト系コンテンツに付与する付与手段とを有する構成を採る。

【0013】本発明のテキスト系コンテンツ管理装置は、メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を受信する受信手段と、受信された追加メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索する検索手段と、検索されたテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得する取得手段と、取得された追加メタデータを送信する送信手段とを有する構成を採る。

【0014】本発明のコンテンツ管理システムは、非テキスト系コンテンツを扱う非テキスト系コンテンツ管理装置と、テキスト系コンテンツを扱うテキスト系コンテンツ管理装置とからなるコンテンツ管理システムであって、前記非テキスト系コンテンツ管理装置は、メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を前記テキスト系コンテンツ管理装置へ送信する第1送信手段と、前記テキスト系コンテンツ管理装置から追加メタデータを受信する第2受信手段と、受信された追加メタデータをメタデータ追加処理対象である非テキスト系コンテンツに付与する付与手段とを有し、前

記テキスト系コンテンツ管理装置は、前記非テキスト系コンテンツ管理装置から追加メタデータ要求を受信する第2受信手段と、受信された追加メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索する検索手段と、検索されたテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得する取得手段と、取得された追加メタデータを前記非テキスト系コンテンツ管理装置へ送信する第2送信手段とを有する構成を採る。

【0015】これらの構成によれば、非テキスト系コンテンツに対して最低限人間の手で付与された意味内容に関するメタデータを元に、より豊富なメタデータを付加することができる。

【0016】本発明の非テキスト系コンテンツ管理装置は、関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を送信する送信手段と、関連コンテンツ付属メタデータを受信する受信手段と、受信された関連コンテンツ付属メタデータを元に、関連コンテンツ情報作成処理対象である非テキスト系コンテンツに対する関連コンテンツ情報を作成する作成手段とを有する構成を採る。

【0017】本発明のテキスト系コンテンツ管理装置は、関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を受信する受信手段と、受信された関連コンテンツ付属メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索する検索手段と、検索されたテキスト系コンテンツと関連するテキスト系コンテンツにあらかじめ付与されたメタデータを関連コンテンツ付属メタデータとして取得する取得手段と、取得された関連コンテンツ付属メタデータを送信する送信手段とを有する構成を採る。

【0018】本発明のコンテンツ管理システムは、非テキスト系コンテンツを扱う非テキスト系コンテンツ管理装置と、テキスト系コンテンツを扱うテキスト系コンテンツ管理装置とからなるコンテンツ管理システムであって、前記非テキスト系コンテンツ管理装置は、関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を前記テキスト系コンテンツ管理装置へ送信する第1送信手段と、前記テキスト系コンテンツ管理装置から関連コンテンツ付属メタデータを受信する第1受信手段と、受信された関連コンテンツ付属メタデータを元に、関連コンテンツ情報作成処理対象である非テキスト系コンテンツに対する関連コンテンツ情報を作成する作成手段とを有し、前記テキスト

系コンテンツ管理装置は、前記非テキスト系コンテンツ管理装置から関連コンテンツ付属メタデータ要求を受信する第2受信手段と、受信された関連コンテンツ付属メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索する検索手段と、検索されたテキスト系コンテンツと関連するテキスト系コンテンツにあらかじめ付与されたメタデータを関連コンテンツ付属メタデータとして取得する取得手段と、取得された関連コンテンツ付属メタデータを前記非テキスト系コンテンツ管理装置へ送信する第2送信手段とを有する構成を採る。

【0019】これらの構成によれば、非テキスト系コンテンツに対して最低限人間の手で付与された意味内容に関するメタデータを元に、非テキスト系コンテンツ同士の内容の関連性を導き出すことができる。

【0020】本発明のメタデータ追加方法は、非テキスト系コンテンツを扱う非テキスト系コンテンツ管理装置と、テキスト系コンテンツを扱うテキスト系コンテンツ管理装置とからなるコンテンツ管理システムにおけるメタデータ追加方法であって、前記非テキスト系コンテンツ管理装置が、メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を前記テキスト系コンテンツ管理装置へ送信するステップと、前記テキスト系コンテンツ管理装置が、前記非テキスト系コンテンツ管理装置から追加メタデータ要求を受信するステップと、前記テキスト系コンテンツ管理装置が、受信した追加メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、前記テキスト系コンテンツ管理装置が、検索したテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得するステップと、前記テキスト系コンテンツ管理装置が、取得した追加メタデータを前記非テキスト系コンテンツ管理装置へ送信するステップと、前記非テキスト系コンテンツ管理装置が、前記テキスト系コンテンツ管理装置から追加メタデータを受信するステップと、前記非テキスト系コンテンツ管理装置が、受信した追加メタデータをメタデータ追加処理対象である非テキスト系コンテンツに付与するステップとを有するようにした。

【0021】この方法によれば、非テキスト系コンテンツに対して最低限人間の手で付与された意味内容に関するメタデータを元に、より豊富なメタデータを付加することができる。

【0022】本発明の関連コンテンツ情報作成方法は、非テキスト系コンテンツを扱う非テキスト系コンテンツ管理装置と、テキスト系コンテンツを扱うテキスト系コンテンツ管理装置とからなるコンテンツ管理システムに

における関連コンテンツ情報作成方法であって、前記非テキスト系コンテンツ管理装置が、関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を前記テキスト系コンテンツ管理装置へ送信するステップと、前記テキスト系コンテンツ管理装置が、前記非テキスト系コンテンツ管理装置から関連コンテンツ付属メタデータ要求を受信するステップと、前記テキスト系コンテンツ管理装置が、受信した関連コンテンツ付属メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、前記テキスト系コンテンツ管理装置が、検索したテキスト系コンテンツと関連するテキスト系コンテンツにあらかじめ付与されたメタデータを関連コンテンツ付属メタデータとして取得するステップと、前記テキスト系コンテンツ管理装置が、取得した関連コンテンツ付属メタデータを前記非テキスト系コンテンツ管理装置へ送信するステップと、前記非テキスト系コンテンツ管理装置が、前記テキスト系コンテンツ管理装置から関連コンテンツ付属メタデータを受信するステップと、前記非テキスト系コンテンツ管理装置が、受信した関連コンテンツ付属メタデータを元に、関連コンテンツ情報作成処理対象である非テキスト系コンテンツに対する関連コンテンツ情報を作成するステップとを有するようにした。

【0023】この方法によれば、非テキスト系コンテンツに対して最低限人間の手で付与された意味内容に関するメタデータを元に、非テキスト系コンテンツ同士の内容の関連性を導き出すことができる。

【0024】本発明のコンテンツ管理プログラムは、メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を送信するステップと、追加メタデータを受信するステップと、受信された追加メタデータをメタデータ追加処理対象である非テキスト系コンテンツに付与するステップとをコンピュータに実行させるためのものである。

【0025】本発明のコンテンツ管理プログラムは、メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を受信するステップと、受信された追加メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、検索されたテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得するステップと、取得された追加メタデータを送信するステップとをコンピュータに実行させるためのものである。

【0026】本発明のコンテンツ管理プログラムは、コンピュータを、非テキスト系コンテンツを扱う非テキスト系コンテンツ管理部およびテキスト系コンテンツを扱うテキスト系コンテンツ管理部として機能させるためのコンテンツ管理プログラムであって、前記コンテンツ管理プログラムは、前記非テキスト系コンテンツ管理部が、メタデータ追加処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む追加メタデータ要求を前記テキスト系コンテンツ管理部へ送信するステップと、前記テキスト系コンテンツ管理部が、前記非テキスト系コンテンツ管理部から追加メタデータ要求を受信するステップと、前記テキスト系コンテンツ管理部が、受信した追加メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、前記テキスト系コンテンツ管理部が、検索したテキスト系コンテンツにあらかじめ付与されたメタデータを追加メタデータとして取得するステップと、前記テキスト系コンテンツ管理部が、取得した追加メタデータを前記非テキスト系コンテンツ管理部へ送信するステップと、前記非テキスト系コンテンツ管理部が、前記テキスト系コンテンツ管理部から追加メタデータを受信するステップと、前記非テキスト系コンテンツ管理部が、受信した追加メタデータをメタデータ追加処理対象である非テキスト系コンテンツに付与するステップとを有するものである。

【0027】これらのプログラムによれば、非テキスト系コンテンツに対して最低限人間の手で付与された意味内容に関するメタデータを元に、より豊富なメタデータを付加することができる。

【0028】本発明のコンテンツ管理プログラムは、関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を送信するステップと、関連コンテンツ付属メタデータを受信するステップと、受信された関連コンテンツ付属メタデータを元に、関連コンテンツ情報作成処理対象である非テキスト系コンテンツに対する関連コンテンツ情報を作成するステップとをコンピュータに実行させるためのものである。

【0029】本発明のコンテンツ管理プログラムは、関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を受信するステップと、受信された関連コンテンツ付属メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、検索されたテキスト系コンテンツと関連するテキスト系コンテンツにあらかじめ付与されたメタデータを関連コンテ

ンツ付属メタデータとして取得するステップと、取得された関連コンテンツ付属メタデータを送信するステップとをコンピュータに実行させるためのものである。

【0030】本発明のコンテンツ管理プログラムは、コンピュータを、非テキスト系コンテンツを扱う非テキスト系コンテンツ管理部およびテキスト系コンテンツを扱うテキスト系コンテンツ管理部として機能させるためのコンテンツ管理プログラムであって、前記コンテンツ管理プログラムは、前記非テキスト系コンテンツ管理部が、関連コンテンツ情報作成処理対象である非テキスト系コンテンツにあらかじめ付与されたオリジナルメタデータを含む関連コンテンツ付属メタデータ要求を前記テキスト系コンテンツ管理部へ送信するステップと、前記テキスト系コンテンツ管理部が、前記非テキスト系コンテンツ管理部から関連コンテンツ付属メタデータ要求を受信するステップと、前記テキスト系コンテンツ管理部が、受信した関連コンテンツ付属メタデータ要求に含まれるオリジナルメタデータを元に、当該オリジナルメタデータに対応する非テキスト系コンテンツと類似するテキスト系コンテンツを検索するステップと、前記テキスト系コンテンツ管理部が、検索したテキスト系コンテンツと関連するテキスト系コンテンツにあらかじめ付与されたメタデータを関連コンテンツ付属メタデータとして取得するステップと、前記テキスト系コンテンツ管理部が、取得した関連コンテンツ付属メタデータを前記非テキスト系コンテンツ管理部へ送信するステップと、前記非テキスト系コンテンツ管理部が、前記テキスト系コンテンツ管理部から関連コンテンツ付属メタデータを受信するステップと、前記非テキスト系コンテンツ管理部が、受信した関連コンテンツ付属メタデータを元に、関連コンテンツ情報作成処理対象である非テキスト系コンテンツに対する関連コンテンツ情報を作成するステップとを有するものである。

【0031】これらのプログラムによれば、非テキスト系コンテンツに対して最低限人間の手で付与された意味内容に関するメタデータを元に、非テキスト系コンテンツ同士の内容の関連性を導き出すことができる。

【0032】

【発明の実施の形態】以下、本発明の実施の形態について、図面を参照して詳細に説明する。なお、本発明はこれら実施の形態に何ら限定されるものではなく、その要旨を逸脱しない範囲において、種々なる態様で実施し得る。

【0033】（実施の形態1）図1は、本発明の実施の形態1におけるコンテンツ管理システムの構成図である。図1に示すコンテンツ管理システムは、非テキスト系コンテンツ管理部110とテキスト系コンテンツ管理部120からなる。

【0034】非テキスト系コンテンツ管理部110は、映像や音声といった非テキスト系コンテンツとそのメ

データを管理するものであって、非テキスト系コンテンツ格納手段111、非テキスト用メタデータ格納手段112、メタデータ入力手段113、要求送信手段114、追加メタデータ取得手段115からなる。

【0035】非テキスト系コンテンツ格納手段111は、非テキスト系コンテンツデータを格納する。

【0036】非テキスト用メタデータ格納手段112は、非テキスト系コンテンツ格納手段111に格納されているコンテンツに対するメタデータを格納する。

【0037】メタデータ入力手段113は、人間の手による非テキスト系コンテンツに対するメタデータの付与を行なう。

【0038】要求送信手段114は、追加メタデータ要求をテキスト系コンテンツ管理部120に対して行なう。

【0039】ここで、追加メタデータ要求とは、非テキスト用メタデータ格納手段112に格納されているメタデータに追加するためのメタデータの要求を意味する。

【0040】追加メタデータ取得手段115は、テキスト系コンテンツ管理部120から送られてきた追加メタデータを取得し、非テキスト用メタデータ格納手段112に格納する。

【0041】テキスト系コンテンツ管理部120は、テキストドキュメントとそのメタデータを管理するもので、テキスト系コンテンツ格納手段121、テキスト用メタデータ格納手段122、メタデータ抽出手段123、要求受信手段124、類似テキスト系コンテンツ検索手段125、追加メタデータ送信手段126からなる。

【0042】テキスト系コンテンツ格納手段121は、テキスト系コンテンツデータを格納する。

【0043】テキスト用メタデータ格納手段122は、テキスト系コンテンツ格納手段121に格納されているコンテンツに対するメタデータを格納する。

【0044】メタデータ抽出手段123は、テキスト系コンテンツ格納手段121に格納されているテキストデータからメタデータを自動的に抽出する。

【0045】要求受信手段124は、非テキスト系コンテンツ管理部110からの追加メタデータ要求を受信する。

【0046】類似テキスト系コンテンツ検索手段125は、追加メタデータ要求のあった非テキスト系コンテンツに類似するテキスト系コンテンツの検索と、当該類似テキスト系コンテンツに付与されたメタデータの取得を行なう。

【0047】追加メタデータ送信手段126は、類似テキスト系コンテンツ検索手段125が取得したメタデータを追加メタデータとして、非テキスト系コンテンツ管理部110に対して送信する。

【0048】次に、本実施の形態におけるメタデータ追

加処理について、具体的な例を用いて説明する。

【0049】図2は、非テキスト系コンテンツ格納手段111に格納されたニュース番組映像およびニュース番組内のニュース項目に付与されたメタデータの一例を示す図である。ニュース番組映像210は更にそのニュースの内容によって複数の項目に分かれる。ここでは、ニュース項目211は、ある野球の試合に関するニュース映像であるとする。このとき、メタデータ220は、ニュース項目211に対してメタデータ入力手段113を通して、人間の手によって付与されたニュース映像の内容に関するメタデータであり、非テキスト用メタデータ格納手段112に格納されている。ここでは、メタデータ220には、ニュースの内容に関する最低限のメタデータが付与されているものとする。また、ニュース項目211には、このニュース項目コンテンツを一意に表わすIDとして、“NEWS_211”が割り当てられているものとする。

【0050】図3は、テキスト系コンテンツ格納手段121に格納された新聞記事とこの新聞記事に付与されたメタデータの一例を示す図である。ここで、新聞記事310は、図2のニュース番組映像210のニュース項目211に類似した内容の記事である。ただし、ニュース項目211と新聞記事310が類似するコンテンツであるという対応付けは、この時点では行なわれていない。また、メタデータ320は、新聞記事310に対して、メタデータ抽出手段123によって自動的に抽出されたメタデータであり、テキスト用メタデータ格納手段122に格納されている。また、新聞記事310には、この新聞記事コンテンツを一意に表わすIDとして、“ARTICLE_310”が割り当てられているものとする。

【0051】なお、テキストデータからのメタデータ抽出方法の一例としては、特開平2001-75959号公報に記載された方法（以下、公知例1と呼ぶ）で実現することが考えられる。図14は、公知例1における文書処理装置の構成を示すブロック図である。公知例1では、入力部1401から入力された文章に対して形態素解析を行なう形態素解析部1402と、形態素列の部分列を重み付きで特定表現候補とする特定表現候補取得部1430と、予めいくつかの特定表現を格納した特定表現辞書1404と、形態素列の特定表現辞書1404中の表現に対するマッチ度を表す実数を、特定表現辞書1404の検索結果として出力する特定表現辞書検索部1405と、特定表現候補に対して、当該特定表現候補に付与された重みと、当該特定表現候補の特定表現辞書1404に対する検索結果とを変数として判別スコアを計算し、当該判別スコアが一定の値を下回る候補を除外する判別分析実行部1406と、特定表現候補のうち、判別分析実行部1406によって除外されなかった形態素の文字列を特定表現として出力する出力部1407とが設けられている。辞書による抽出とパタンマッチングに

よる抽出とをバランス良く組み合わせて、人名等を的確に抽出することが可能となる。また、テキストデータからのメタデータ抽出方法は、上記公知例1以外にも既に述べたように様々な研究がなされており、ここではその方法を特に限定することはない。また、自動抽出されたメタデータ320は、人間の手によって付与されたメタデータ220と比較して、ニュース本題に関するメタデータのみならず、その前後関係やより詳細なキーワードを含むものとなっている。

10 【0052】なお、図2、図3において、メタデータはXML（eXtensible Markup Language）形式で記述されているが、これはメタデータの記述形式の一例であり、その他どのような記述形式であっても構わない。また、メタデータとして複数のキーワードを併記する形式をとっているが、各キーワードに5W1Hのような意味を持たせたり、もしくはフリーテキスト形式のメタデータを持たせたりしても構わない。

【0053】図4は、実施の形態1における非テキスト系コンテンツに対するメタデータ追加処理フローを示す図である。以下、例えば、図2に示すニュース項目211に対するメタデータを追加する処理について、図4に沿って説明する。

【0054】ステップ401：非テキスト系コンテンツ管理部110の要求送信手段114にて、メタデータ追加処理対象の非テキスト系コンテンツに付与されているメタデータを、非テキスト用メタデータ格納手段112から取得し、この取得したメタデータ（以下、オリジナルメタデータと呼ぶ）とともに追加メタデータ要求をテキスト系コンテンツ管理部120へ送信する。この例の場合、オリジナルメタデータとしてメタデータ220を取得し、追加メタデータ要求とともに送信する。

【0055】ステップ402：テキスト系コンテンツ管理部120の要求受信手段124にて、非テキスト系コンテンツ管理部110からの追加メタデータ要求（オリジナルメタデータを含む）を受信する。

【0056】ステップ403：類似テキスト系コンテンツ検索手段125にて、前記追加メタデータ要求に含まれていたオリジナルメタデータを利用して、類似するテキスト系コンテンツを検索し、当該類似テキスト系コンテンツに付与されているメタデータをテキスト用メタデータ格納手段122から取得する。ここで、類似するテキスト系コンテンツが複数検索された場合には、その類似度が最も高いテキスト系コンテンツに付与されているメタデータを取得する。また、「類似」とは、非テキスト系コンテンツとテキスト系コンテンツとの情報の重複度合が所定の要件を満たす場合をいう。

【0057】例えば、キーワードを利用した情報検索方法の一例としては、特開平10-49549号公報に記載された方法（以下、公知例2と呼ぶ）で実現することが考えられる。図15は、公知例2における文書検索装

置の構成を示したブロック図である。公知例2では、頻度スコア算定手段1508にて、単語頻度算定手段1507が出力した総文書数、単語の出現する文書数、文書における単語の出現頻度、単語の重み付けパラメータから、単語頻度による文書と検索要求の合致度合を示す頻度スコアを算出し、文書スコア算定手段1509にて、上記頻度スコアから文書と検索要求の合致度合を示す文書スコア算出し、順位付けを行なうことによって、検索意図により近い検索結果を得ることが可能となる。また、メタデータ（キーワード）を利用した情報検索方法は、上記公知例2以外にも既に述べたアメリカのTipsterプロジェクトやSIGIR (Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, July 24-28, 2000を参照)などで様々な研究がなされており、ここではその方法を特に限定することはしない。この例の場合、類似するテキスト系コンテンツ検索結果として新聞記事310が導き出されたとすると、この新聞記事310に付与されているメタデータ320を取得することになる。

【0058】ステップ404：追加メタデータ送信手段126にて、前記類似テキスト系コンテンツ検索手段125が取得したメタデータを追加メタデータとして、非テキスト系コンテンツ管理部110に送信する。

【0059】ステップ405：非テキスト系コンテンツ管理部110の追加メタデータ取得手段115にて、テキスト系コンテンツ管理部120からの追加メタデータを受信し、この追加メタデータをメタデータ追加処理対象の非テキスト系コンテンツに付与された追加メタデータとして、非テキスト用メタデータ格納手段112に格納する。

【0060】なお、上記ステップ401から405のステップに記載した非テキスト系コンテンツ管理部110とテキスト系コンテンツ管理部120の処理は、コンピュータに上記ステップの処理を実行させるプログラムをインストールすることにより、実現することが可能である。

【0061】図5は、ニュース項目211に対して追加メタデータ取得手段115が実行したメタデータ追加処理が終わった段階で、付与されているメタデータの例を示す図である。メタデータ501の例では、追加メタデータ取得手段115が受信した追加メタデータをそのまま付与している。一方、メタデータ502の例では、追加メタデータ取得手段115が受信した追加メタデータとオリジナルメタデータを比較し、重複したメタデータを除いて付与している。本実施の形態では、上記2種類の方法のうちどちらの方法をとっても構わない。

【0062】図6は、実施の形態1における非テキスト系コンテンツに対するメタデータ追加処理でのコンテンツとメタデータの関係の集約図である。この図は、メタ

データ付与対象の非テキスト系コンテンツに対する類似テキスト系コンテンツが求められた場合に、当該類似テキスト系コンテンツに対して抽出されている豊富なメタデータを、メタデータ付与対象の非テキスト系コンテンツに対して追加することによって、非テキスト系コンテンツに対するメタデータを豊富にしていることを表している。

【0063】以上のように、本実施の形態では、メタデータ追加処理対象の非テキスト系コンテンツに対して人間の手によって付与されたメタデータを利用して、類似するテキスト系コンテンツを検索し、当該類似テキスト系コンテンツに対して自動的に抽出されたメタデータを、メタデータ追加処理対象の非テキスト系コンテンツに対する追加メタデータとして取得することにより、非テキスト系コンテンツに対するメタデータを、人間の手によって付与された、限られた数から増やすことが可能となる。

【0064】更に、このようにコンテンツに対するメタデータを豊富にすることによって、メタデータを利用した非テキスト系コンテンツの検索において当該コンテンツの再現率が高まるといった二次的な効果が得られる。

【0065】なお、本実施の形態においては、テキスト系コンテンツの例として図3に示したようにテキストのみの新聞記事を挙げて説明したが、例えば、図や写真を含んだHTML形式のドキュメントであっても構わない。

【0066】また、本実施の形態においては、非テキスト系コンテンツ管理部110とテキスト系コンテンツ管理部120は、2つが同一コンピュータ上に存在して、両者から成る1つのコンテンツ管理装置としても実現可能であり、また、別々のコンピュータ上に存在して、情報伝送可能なネットワークを介して構成されるコンテンツ管理システムとしても実現可能である。

【0067】また、本実施の形態においては、非テキスト系コンテンツ管理部110とテキスト系コンテンツ管理部120が1対1に存在する構成として説明したが、1つの非テキスト系コンテンツ管理部が、複数のテキスト系コンテンツ管理サーバに対して追加メタデータ要求を送信するような1対nの構成であっても構わない。

【0068】また、本実施の形態の図2においては、ニュース番組映像210というコンテンツの一部分であるニュース項目211に対してメタデータを付与する場合を例にとって説明したが、メタデータを割り当てる対象は、コンテンツ全体に対してであっても、コンテンツの一部に対してであっても構わない。

【0069】また、本実施の形態の図4のステップ403においては、類似するテキスト系コンテンツが複数検索された場合には、その類似度が最も高いテキスト系コンテンツに付与されているメタデータを取得するものとして説明したが、それ以外にも、例えば、類似度の高い

方から複数件分（例えば10件分）のメタデータを取得し、これら複数件分のメタデータを追加メタデータとしてステップ405において非テキスト用メタデータ格納手段112へ格納するようにしても構わない。

【0070】また、類似テキスト系コンテンツ検索手段125における検索処理、および、メタデータ抽出手段123におけるメタデータ抽出処理においては、その処理を完全に自動化するのではなく、得られた結果を人間がチェックして取捨選択することにより精度を上げるような方法をとっても構わない。

【0071】（実施の形態2）以下、本発明の実施の形態2について説明する。本実施の形態におけるコンテンツ管理システムの構成図は、図7に示すように、実施の形態1の構成図（図1）から追加メタデータ取得手段115および追加メタデータ送信手段126を除くとともに、新たに関連コンテンツ付属メタデータ取得手段711、類似非テキスト系コンテンツ検索手段712、非テキスト用関連コンテンツ情報格納手段713、関連コンテンツ情報作成手段721、テキスト用関連コンテンツ情報格納手段722、および関連コンテンツ付属メタデータ送信手段723を備えている。

【0072】関連コンテンツ付属メタデータ取得手段711は、テキスト系コンテンツ管理部120aから送られてきた関連コンテンツ付属メタデータを取得する。

【0073】類似非テキスト系コンテンツ検索手段712は、関連コンテンツ付属メタデータを元に、非テキスト系コンテンツに対する関連コンテンツ情報を作成する。

【0074】非テキスト用関連コンテンツ情報格納手段713は、非テキスト系コンテンツ格納手段111に格納されているコンテンツ間の関連情報を格納する。

【0075】関連コンテンツ情報作成手段721は、テキスト用メタデータ格納手段122に格納されているメタデータを元に、テキスト系コンテンツ格納手段121に格納された各コンテンツ間の関連情報を自動的に作成する。

【0076】テキスト用関連コンテンツ情報格納手段722は、前記関連コンテンツ情報作成手段721において作成された関連コンテンツ情報を格納する。

【0077】関連コンテンツ付属メタデータ送信手段723は、類似テキスト系コンテンツ検索手段125が取得した類似テキストコンテンツに対する関連コンテンツ群に付与されているメタデータ群を、非テキスト系コンテンツ管理部110aに対して送信する。

【0078】次に、本実施の形態における関連コンテンツ情報作成処理について、具体的な例を用いて説明する。

【0079】図8は、図2と同様に非テキスト系コンテンツ格納手段111に格納されたニュース番組映像およびニュース番組内のニュース項目に付与されたメタデー

タの別の一例を示す図である。ここでは、ニュース項目813も、ある野球の試合に関するニュース映像であり、メタデータ820は、ニュース項目813に対して人間の手によって付与されたニュース映像の内容に関するメタデータであるとする。また、ニュース項目813には、このニュース項目コンテンツを一意に表わすIDとして、“NEWS_813”が割り当てられているものとする。

【0080】図9は、図3と同様にテキスト系コンテンツ格納手段121に格納された新聞記事およびこの新聞記事に付与されたメタデータの別の一例を示す図である。ここで、新聞記事910は、図8のニュース映像810のニュース項目813に類似した内容の記事である。また、メタデータ920は、新聞記事810に対して、メタデータ抽出手段123によって自動的に抽出されたメタデータであり、テキスト用メタデータ格納手段122に格納されている。また、新聞記事910には、この新聞記事コンテンツを一意に表わすIDとして、“ARTICLE_910”が割り当てられているものとする。

【0081】図10は、関連コンテンツ情報作成手段721によって自動的に作成された関連コンテンツ情報の一例を示す図である。例えば、図10の関連コンテンツ情報1001の場合、コンテンツのIDが“ARTICLE_310”のコンテンツの関連記事として、IDが“ARTICLE_910”のコンテンツが挙げられている。なお、テキストデータの関連性を検出するようなテキスト処理に関する技術は、基本的に実施の形態1のステップ403で説明したように、類似コンテンツを見つけるためのキーワードを利用した情報検索方法と同様の技術である。本明細書においては、非テキスト系コンテンツとテキスト系コンテンツとの情報の重複度が所定の要件を満たす場合には「類似」、テキスト系コンテンツ同士または非テキスト系コンテンツ同士が所定の手法により関係付けられる場合には「関連」という語を用いて区別している。

【0082】また、テキスト系コンテンツの場合、あらかじめコンテンツ同士の関連情報（続報記事やリンク）を持っている場合も多いので、これらの情報を元に関連コンテンツ情報を作成しても構わない。

【0083】また、図10の関連コンテンツ情報1002に示したように、ある1つのコンテンツに対して、複数の関連コンテンツを持つような関連コンテンツ情報を作成することも可能である。

【0084】図11は、実施の形態2における非テキスト系コンテンツに対する関連コンテンツ情報作成処理フローを示す図である。以下、例えば、図2に示すニュース項目211に対する関連コンテンツ情報を作成する処理について、図11に沿って説明する。

【0085】ステップ1101：非テキスト系コンテンツ管理部110aの要求送信手段114にて、関連コンテンツ情報作成処理対象の非テキスト系コンテンツに付与されているメタデータを、非テキスト用メタデータ格

10

20

30

40

50

納手段112から取得し、この取得したオリジナルメタデータとともに関連コンテンツ付属メタデータ要求をテキスト系コンテンツ管理部120aへ送信する。この例の場合、オリジナルメタデータとしてメタデータ220を取得し、関連コンテンツ付属メタデータ要求とともに送信する。

【0086】ここで、関連コンテンツ付属メタデータ要求とは、一般には、一的非テキスト系コンテンツデータに関連する他の非テキスト系コンテンツを得るために必要なメタデータの要求を示すが、本実施の形態では、非

テキスト系コンテンツ格納手段111に格納された一的非テキスト系コンテンツに関連する他の非テキスト系コンテンツデータを得るために、当該他の非テキスト系コンテンツと類似するテキスト系コンテンツデータに付与されたメタデータの要求を示すものとする。

【0087】ステップ1102：テキスト系コンテンツ管理部120aの要求受信手段124にて、非テキスト系コンテンツ管理部110aからの関連コンテンツ付属メタデータ要求（オリジナルメタデータを含む）を受信する。

【0088】ステップ1103：類似テキスト系コンテンツ検索手段125にて、前記関連コンテンツ付属メタデータ要求に含まれていたオリジナルメタデータを利用して、類似するテキスト系コンテンツを検索し、当該類似テキスト系コンテンツのコンテンツIDを取得する。ここで、類似するテキスト系コンテンツが複数検索された場合には、その類似度が最も高いテキスト系コンテンツに付与されているメタデータを取得する。この例の場合、類似するテキスト系コンテンツ検索結果として、新聞記事310が導き出されたとすると、コンテンツID"ARTICLE_310"を取得することになる。

【0089】ステップ1104：関連コンテンツ付属メタデータ送信手段723にて、テキスト用関連コンテンツ情報格納手段722に格納された情報を参照して、前記類似テキスト系コンテンツ検索手段125が取得したコンテンツIDの関連コンテンツIDを取得する。この例の場合、図10の関連コンテンツ情報1001からわかるように、"ARTICLE_910"を取得する。

【0090】ステップ1105：更に関連コンテンツ付属メタデータ送信手段723にて、前記ステップ1104で取得した関連コンテンツIDで特定されるテキスト系コンテンツに付与されているメタデータをテキスト用メタデータ格納手段122から取得し、これを関連コンテンツ付属メタデータとして、非テキスト系コンテンツ管理部110aに送信する。この例の場合、コンテンツID"ARTICLE_910"で特定される新聞記事910に付与されているメタデータ920を送信することになる。

【0091】ステップ1106：非テキスト系コンテンツ管理部110aの関連コンテンツ付属メタデータ取得手段711にて、テキスト系コンテンツ管理部120a

からの関連コンテンツ付属メタデータを受信する。

【0092】ステップ1107：類似非テキスト系コンテンツ検索手段712にて、前記関連コンテンツ付属メタデータ取得手段711が取得した関連コンテンツ付属メタデータを利用して、類似する非テキスト系コンテンツを検索し、当該類似非テキスト系コンテンツのコンテンツIDを取得する。ここで、類似する非テキスト系コンテンツが複数検索された場合には、その類似度が最も高い非テキスト系コンテンツのコンテンツIDを取得する。この例の場合、類似する非テキスト系コンテンツ検索結果として、図8のニュース項目813が導き出されたとすると、コンテンツID"NEWS_813"を取得することになる。

【0093】ステップ1108：更に類似非テキスト系コンテンツ検索手段712にて、前記ステップ1107で取得したコンテンツIDと、関連コンテンツ情報作成処理対象コンテンツのコンテンツIDとで、関連コンテンツ情報を作成し、非テキスト用関連コンテンツ情報格納手段713へ格納する。

【0094】なお、本実施の形態の図11のステップ1103においては、類似するテキスト系コンテンツが複数検索された場合には、その類似度が最も高いテキスト系コンテンツに付与されているメタデータを取得するものとして説明したが、それ以外にも、例えば、類似度の高い方から複数件分（例えば10件分）のメタデータを取得するようにしても構わない。

【0095】また、上記ステップ1104でコンテンツID"ARTICLE_910"によって特定されるテキスト系コンテンツに付与されたメタデータを送信する代わりに、ステップ1103で得たコンテンツID"ARTICLE_310"によって特定されるテキスト系コンテンツに付与されたメタデータを関連コンテンツ付属メタデータ送信手段723にて、非テキスト系コンテンツ管理部110aへ送信するようにしても良い。この場合、類似非テキスト系コンテンツ検索手段712は、上記コンテンツID"ARTICLE_310"によって特定されるテキスト系コンテンツに付与されたメタデータに類似するメタデータを有する非テキスト系コンテンツを検索することとなる。

【0096】また、上記ステップ1104で得たコンテンツID"ARTICLE_910"と更に関連するコンテンツIDを検索するといったように、連鎖的に検索を行なうことも可能である。

【0097】また、上記ステップ1104において関連コンテンツIDが複数存在する場合には、上記ステップ1105にて前記複数の関連コンテンツIDそれぞれに対応する関連コンテンツ付属メタデータを取得し、上記ステップ1107にて前記複数の関連コンテンツ付属メタデータごとに、類似する非テキスト系コンテンツを検索してそれぞれ当該類似非テキスト系コンテンツのコンテンツIDを取得し、上記ステップ1108にて前記ス

ステップ1107で複数取得したコンテンツID群と関連コンテンツ情報作成処理対象コンテンツのコンテンツIDとで関連コンテンツ情報を作成することとなる。

【0098】図12には、上記の例においてニュース項目211に対する関連コンテンツ情報作成処理が終わった段階で非テキスト用関連コンテンツ情報格納手段713へ格納された関連コンテンツ情報の一例が示されている。

【0099】図13は、実施の形態2における非テキスト系コンテンツに対する関連コンテンツ情報作成処理でのコンテンツとメタデータの関係の集約図である。例えば、図13に示すような人間の手によって付与された最低限のメタデータ220とメタデータ820だけでは、ニュース項目211とニュース項目813が関連するコンテンツであるとは判断できない。しかし、前記2つの非テキスト系コンテンツと類似性のあるテキスト系コンテンツである記事310と記事910の関連情報を非テキスト系コンテンツ側に写像することによって、前記2つの非テキスト系コンテンツが、同じ「5月21日」に行なわれた「AチームvsBチーム」の試合に関する関連ニュース項目であることが導き出される。すなわち、図11に示されたステップを実行することにより、ニュース項目211とニュース項目813が関連コンテンツであることが導き出される。

【0100】以上のように、本実施の形態では、関連コンテンツ情報作成処理対象の非テキスト系コンテンツに対して人間の手によって付与されたメタデータを利用して、類似するテキスト系コンテンツを検索し、当該類似テキスト系コンテンツに対して関連付けがあらかじめされているテキスト系コンテンツに対して自動的に抽出されたメタデータ（関連コンテンツ付属メタデータ）を利用して、類似する非テキスト系コンテンツを検索することにより、最低限人間の手によって付与されたメタデータのみでは導き出すことができなかった非テキスト系コンテンツ同士の関連性を導き出すことが可能となる。

【0101】また、本実施の形態においても、実施の形態1と同様に、非テキスト系コンテンツ管理部110aとテキスト系コンテンツ管理部120aは、2つが同一コンピュータ上に存在して、両者から成る1つのコンテンツ管理装置としても実現可能であり、また、別々のコンピュータ上に存在して、情報伝送可能なネットワークを介して構成されるコンテンツ管理システムとしても実現可能である。

【0102】また、上記ステップ1101から1108のステップに記載した非テキスト系コンテンツ管理部110aとテキスト系コンテンツ管理部120aの処理は、コンピュータに上記ステップの処理を実行させるプログラムをインストールすることにより、実現することが可能である。

【0103】

【発明の効果】以上説明したように、本発明によれば、メタデータ追加処理対象の非テキスト系コンテンツに対して人間の手によって付与されたメタデータを利用して、類似するテキスト系コンテンツを検索し、当該類似テキスト系コンテンツに対して自動的に抽出されたメタデータを、メタデータ追加処理対象の非テキスト系コンテンツに対する追加メタデータとして取得することにより、MPEG-7がメタデータ付与の対象とするような非テキスト系コンテンツに対するメタデータを、人間の手によって付与された、限られた数から増やすことが可能となる。

【0104】更に、このようにコンテンツに対するメタデータを豊富にすることによって、メタデータを利用した非テキスト系コンテンツの検索において当該コンテンツの再現率が高まるといった二次的な効果が得られる。

【0105】また、関連コンテンツ情報作成処理対象の非テキスト系コンテンツに対して人間の手によって付与されたメタデータを利用して、類似するテキスト系コンテンツを検索し、当該類似テキスト系コンテンツに対して関連付けがあらかじめされているテキスト系コンテンツに対して自動的に抽出されたメタデータ（関連コンテンツ付属メタデータ）を利用して、類似する非テキスト系コンテンツを検索することにより、最低限人間の手によって付与されたメタデータのみでは導き出すことができなかった非テキスト系コンテンツ同士の関連性を導き出すことが可能となる。

【図面の簡単な説明】

【図1】本発明の実施の形態1におけるコンテンツ管理システムの構成図

【図2】本発明の実施の形態1における非テキスト系コンテンツとそのメタデータの一例を示す図

【図3】本発明の実施の形態1におけるテキスト系コンテンツとそのメタデータの一例を示す図

【図4】本発明の実施の形態1における非テキスト系コンテンツに対するメタデータ追加処理フローを示す図

【図5】本発明の実施の形態1におけるニュース項目211に対してメタデータ追加処理が終わった段階で、付与されているメタデータの例を示す図

【図6】本発明の実施の形態1における非テキスト系コンテンツに対するメタデータ追加処理でのコンテンツとメタデータの関係の集約図

【図7】本発明の実施の形態2におけるコンテンツ管理システムの構成図

【図8】本発明の実施の形態2における非テキスト系コンテンツとそのメタデータの一例を示す図

【図9】本発明の実施の形態2におけるテキスト系コンテンツとそのメタデータの一例を示す図

【図10】本発明の実施の形態2における関連コンテンツ情報作成手段721によって自動的に作成された関連コンテンツ情報の一例を示す図

【図11】本発明の実施の形態2における非テキスト系コンテンツに対する関連コンテンツ情報作成処理フローを示す図

【図12】本発明の実施の形態2における非テキスト用関連コンテンツ情報格納手段713へ格納されたコンテンツ関連情報の一例を示す図

【図13】本発明の実施の形態2における非テキスト系コンテンツに対する関連コンテンツ情報作成処理でのコンテンツとメタデータの関係の集約図

【図14】公知例1における文書処理装置の構成を示すブロック図

【図15】公知例2における文書検索装置の構成を示したブロック図

【符号の説明】

110、110a 非テキスト系コンテンツ管理部

111 非テキスト系コンテンツ格納手段

112 非テキスト用メタデータ格納手段

113 メタデータ入力手段

114 要求送信手段

115 追加メタデータ取得手段

120、120a テキスト系コンテンツ管理部

121 テキスト系コンテンツ格納手段

122 テキスト用メタデータ格納手段

123 メタデータ抽出手段

124 要求受信手段

125 類似テキスト系コンテンツ検索手段

126 追加メタデータ送信手段

711 関連コンテンツ付属メタデータ取得手段

712 類似非テキスト系コンテンツ検索手段

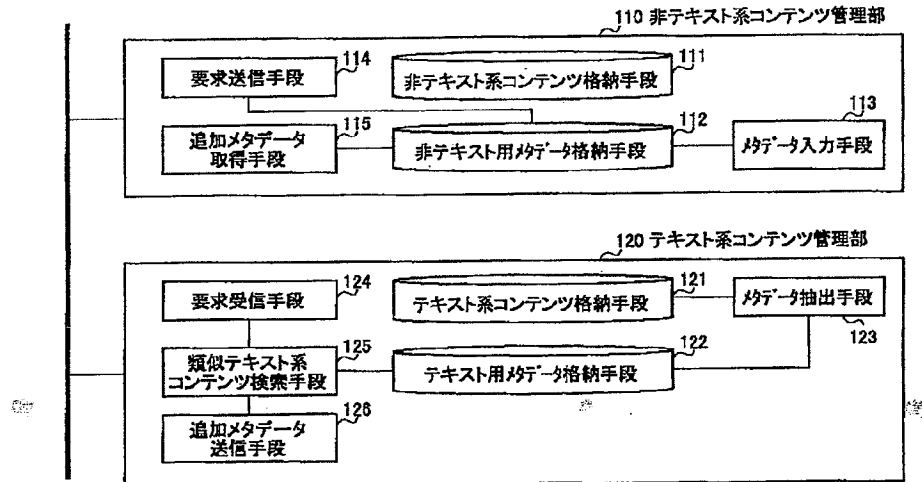
713 非テキスト用関連コンテンツ情報格納手段

721 関連コンテンツ情報作成手段

722 テキスト用関連コンテンツ情報格納手段

723 関連コンテンツ付属メタデータ送信手段

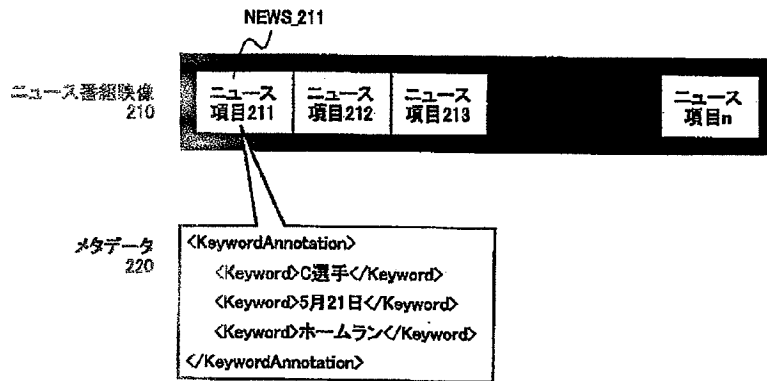
【図1】



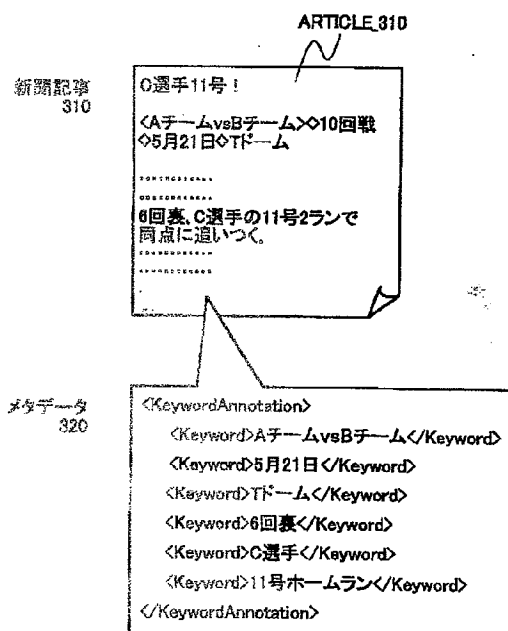
【図10】

関連コンテンツ情報	コンテンツID	
	コンテンツID	関連コンテンツID
1001	ARTICLE_310	ARTICLE_910
1002	ARTICLE_311	ARTICLE_400, ARTICLE_530, ARTICLE_683
	⋮	⋮

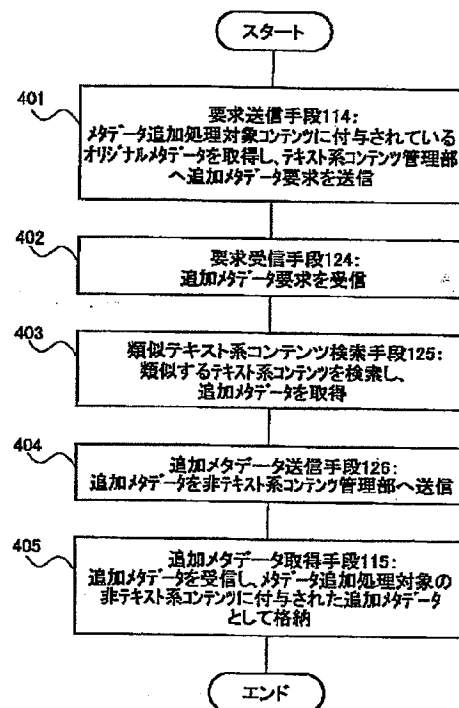
【図2】



【図3】



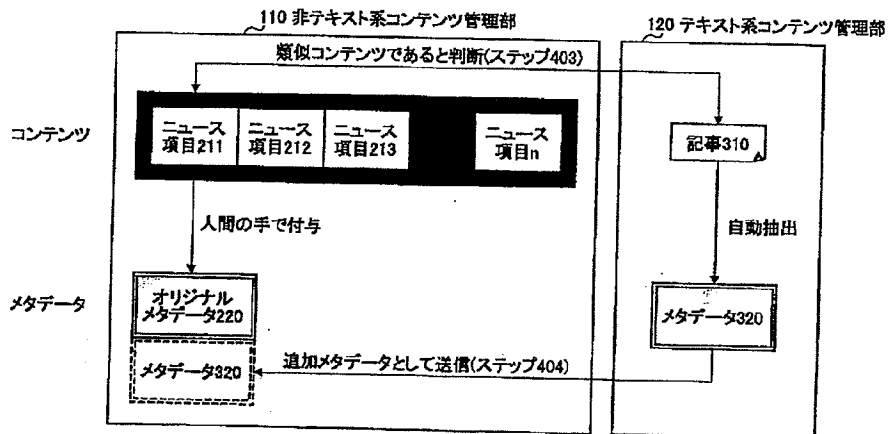
【図4】



【図5】

	オリジナルメタデータ	追加メタデータ
メタデータ 501	<KeywordAnnotation> <Keyword>〇選手</Keyword> <Keyword>5月21日</Keyword> <Keyword>ホームラン</Keyword> </KeywordAnnotation>	<KeywordAnnotation> <Keyword>AチームvsBチーム</Keyword> <Keyword>5月21日</Keyword> <Keyword>Tチーム</Keyword> <Keyword>8回裏</Keyword> <Keyword>〇選手</Keyword> <Keyword>11号ホームラン</Keyword> </KeywordAnnotation>
メタデータ 502	<KeywordAnnotation> <Keyword>〇選手</Keyword> <Keyword>5月21日</Keyword> <Keyword>ホームラン</Keyword> </KeywordAnnotation>	<KeywordAnnotation> <Keyword>Tチーム</Keyword> <Keyword>8回裏</Keyword> <Keyword>11号ホームラン</Keyword> </KeywordAnnotation>

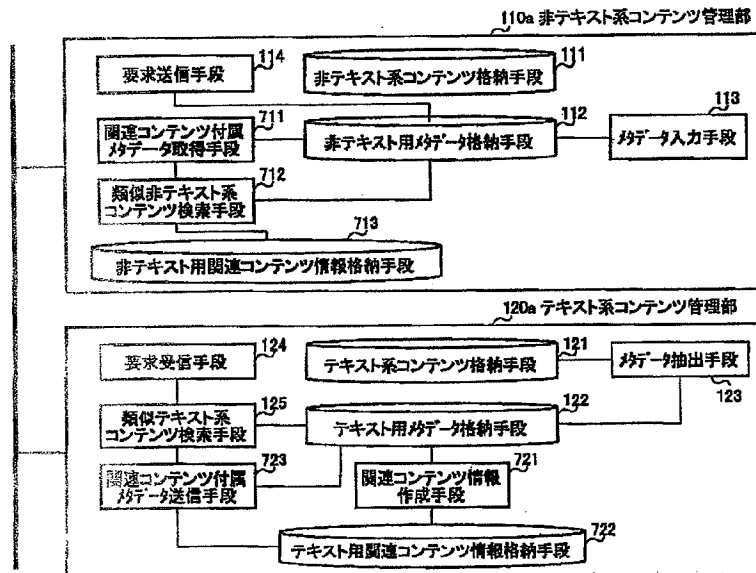
【図6】



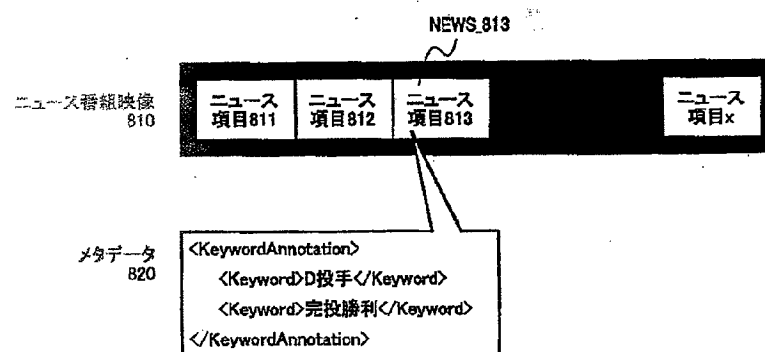
【図12】

関連コンテンツ情報 1201	コンテンツID	関連コンテンツID
	NEWS_211	NEWS_813

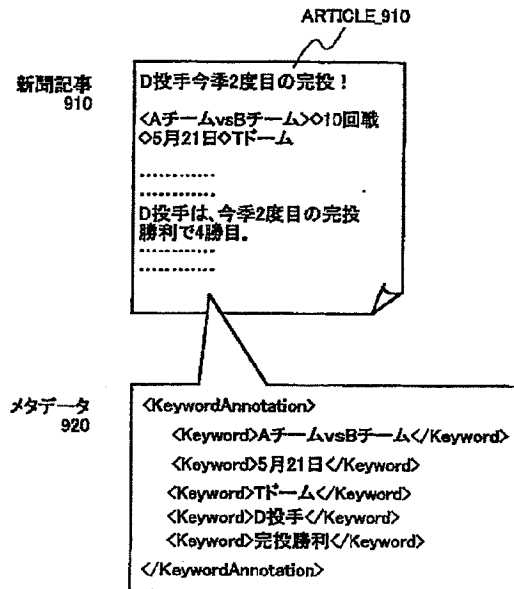
【図7】



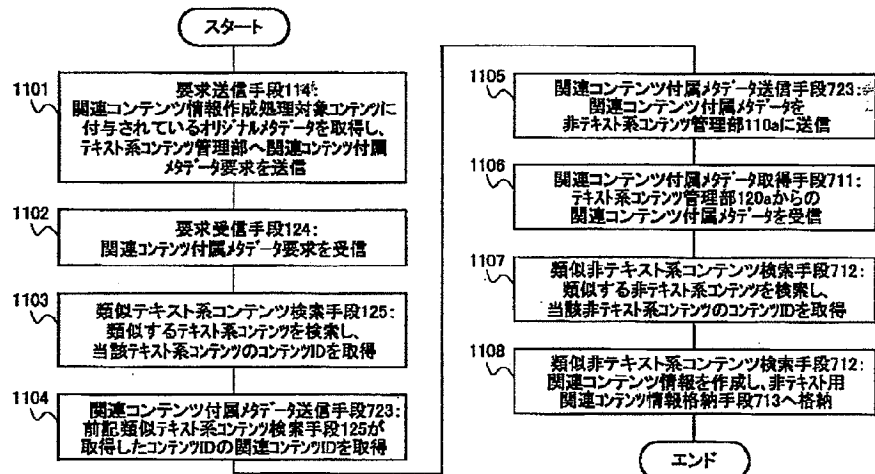
【図8】



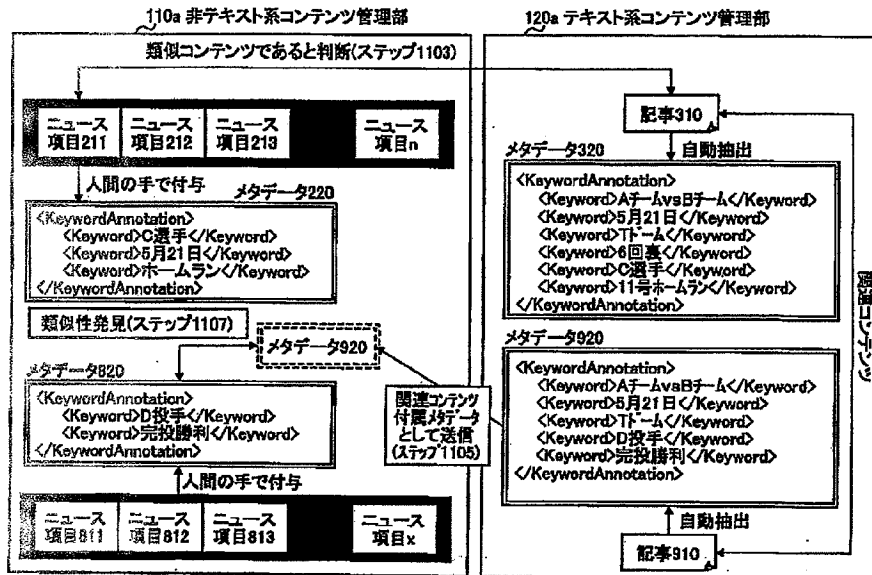
【図9】



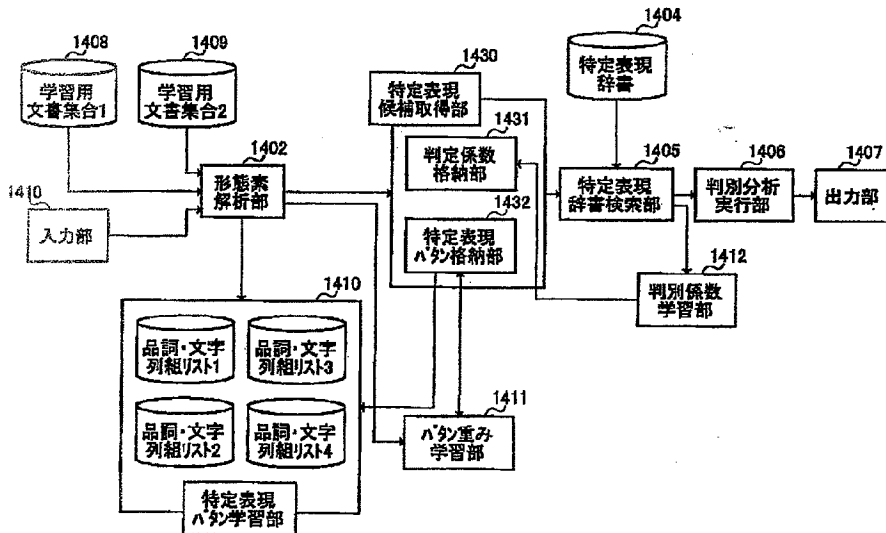
【図11】



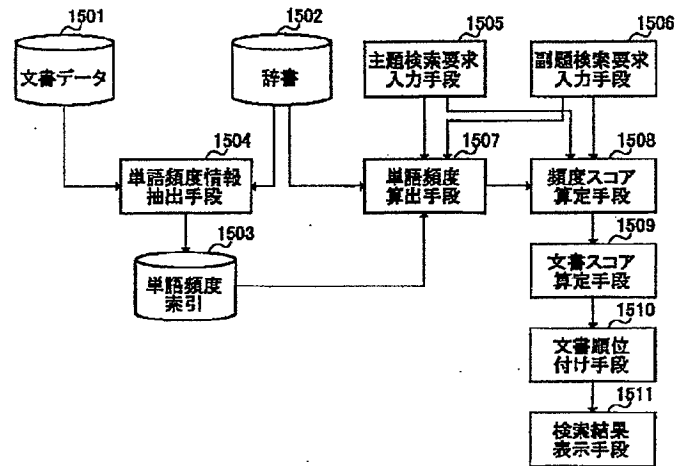
【図13】



【図14】



【図15】



フロントページの続き

Fターム(参考) 5B075 ND03 ND16 NK02 NR05 NR14

PP22 PR06 QM08

5B082 AA13 EA07